2023-08-25

National Cloud Strategy 2023

Alliance National Cloud Strategy Working Group
Version 1

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Table of Contents

Definitions

Term	Definition
ARC (Advanced Research Computing)	The elements required to perform computational and data-intensive research and data-management, including high-performance computing and storage. ARC relies on high-speed networks, software, standards, and data management services.
Capital Expenditure (CAPEX)	The money an organization or corporate entity spends to buy, maintain, or improve its fixed assets.[1]
Cloud Computing	Cloud computing is "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction"[2].
Community Cloud	Cloud resources and services operated by members of a community for the collective benefit of that community.
Commercial Cloud	Cloud services operated as a profit-making business, available to anyone who can pay the cost.
Cloud Service Providers (CSPs)	A cloud service provider, or CSP, is a company that offers some components of cloud computing to other businesses or individuals.
Cloud Infrastructure	The physical and logical components, systems, and services on which cloud services are built.
Cloud Service	There are 3 main classes of cloud service offerings: Infrastructure-as-a-Service (IaaS), Platforms-as-a-Service (PaaS), and Software-

[1] Jason Fernando, "Capital Expenditure (CapEx) Definition, Formula, and Examples," Investopedia, May, 2023, https://www.investopedia.com/terms/c/capitalexpenditure.asp.

[2] Peter Mell and Timothy Grance, *The NIST Definition of Cloud Computing* (Gaithersburg: National Institute of Standards and Technology, 2011) https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf.

Offerings	as-a-Service (SaaS)
Containers	Containers are used to abstract applications from the physical environment in which they are running. A container packages all dependencies related to a software component, and runs them in an isolated environment.
High Performance Computing (HPC)	High performance computing (HPC) is the ability to process data and perform complex calculations at high speeds and usually implies low latency networking (interconnect) between nodes, such as Infiniband or OmniPath.
Hybrid Cloud	Hybrid cloud refers to a mixed computing, storage, and services environment made up of on-premises infrastructure, community cloud services, and a commercial cloud with orchestration among the various platforms. NIST Definition[3]: *The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or commercial) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).*
Infrastructure as a Service (IaaS)	Infrastructure as a service (IaaS) is a type of cloud computing service that offers essential compute, storage, and networking resources on demand. IaaS is one of the three classes of cloud services, along with software as a service (SaaS), platform as a service (PaaS).
NIST (National Institute of Standards and Technology)	The world's leader in creating critical measurement solutions and promoting equitable standards. The National Institute of Standards and Technology is a physical sciences laboratory and non-regulatory agency of the United States Department of Commerce.
Operating Expenses (OPEX)	The money an organization or corporate entity spends to use assets for a fixed period.
Platform as a Service (PaaS)	Platform as a service (PaaS) is a cloud computing model where a third-party provider delivers application development and deployment tools to end users, typically hosted on the provider's own IaaS cloud resources.

[3] Mell and Grance, *The NIST Definition*.

Private Cloud	Cloud services operated by a single organization exclusively for its own use. NIST Definition[4]: *The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.*
Public Cloud	A term sometimes used as a synonym for Commercial Cloud. NIST Definition[5]: *The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.*
Software as a Service (SaaS)	Software as a service (SaaS) is a software distribution model in which a cloud provider hosts entire applications, including the platform and infrastructure resources required to run them, and makes the applications available directly to end users over the internet.

Introduction

The Digital Research Alliance of Canada exists to serve Canadian researchers – ultimately propelling Canada onto the international stage as a leader in the knowledge economy. Its coordination and funding of activities in advanced research computing (ARC), research data management (RDM) and research software (RS) will benefit Canadian researchers. This document establishes and communicates the organisation's unified cloud computing adoption strategy and direction. It identifies high-level approaches and methodologies upon which all impacted stakeholders agree.

As Canadian researchers leverage DRI resources, the Alliance is well-positioned to facilitate solutions to Canadian research groups' challenges in accessing appropriate DRI tools, resources, technology services and capabilities. As the Alliance begins to address DRI's gaps, it

[4] Mell and Grance, *The NIST Definition*.

[5] Mell and Grance, *The NIST Definition*.

will be essential to continue supporting defined research projects through services such as the Community Clouds, Resource Allocations Competition (RAC), Commercial Cloud offerings, and research software development. These services will enhance the community's understanding of DRI through research activities and foster the expansion of the Alliance's DRI ecosystem.

Background

In 2021, the Alliance kicked off a process to develop a National Cloud Strategy, creating a Working Group of cloud experts[6] from the Alliance and the Federation. The initial draft strategy was never published, but the working group did make a number of recommendations:

1. *To grow the foundation of the Community Cloud resources and support at a scale aligned with the demand from Canadian researchers.*
2. *To create, staff, and fund a dedicated Cloud Researcher Support group with a mandate to advise Canadian researchers on cloud technologies best adapted to their specific needs and provide subsequent technical support.*
3. *Identify applicable Commercial Cloud capabilities by partnering the Alliance's Cloud Support group with engaged research projects in pilot programs.*
4. *Develop an approach to providing Commercial and other cloud services and resources to Canadian researchers using competitive calls and the acquisition/deployment of Commercial Cloud services that can be allocated to researchers on a strategic basis.*
5. *Identify common research use cases and build/enhance research software platforms addressing them effectively through a defined needs assessment and investment strategy process.*
6. *Develop a set of use cases that demonstrate the value of commercial and community cloud in the research landscape, providing a framework for a more detailed cloud/hybrid cloud strategy for the next 3-10 years.*

A substantial response to these recommendations (especially 1, 2, and 4) were reflected in the Alliance's Multi-Year Funding Proposal for 2023-25 (MYFP25). Key among those initiatives that addressed those recommendations are:[7]

1. **ARC Infrastructure Renewal and Commercial Cloud Pilot**: *Addresses the critical issue of Canada's aging ARC infrastructure (HPC Renewal and Community Cloud Renewal) across the National Host Sites. In addition, ISED has directed the Alliance to*

[6] See Appendix A.

[7] "ISED approves funding for digital research infrastructure (DRI) initiatives," Digital Research Alliance of Canada, March 28, 2023, https://alliancecan.ca/en/latest/news/ised-approves-funding-digital-research-infrastructure-dri-initiatives.

develop a commercial cloud strategy and pilot proposal to complement existing community cloud solutions.

2. **Data Centre Upgrade**: *Upgrades the existing power infrastructure to 5 MW with related cooling capacity at the McGill/CQ national site, allowing future HPC and Cloud expansions on this site.*
3. **Extending User Support for 2023-2025**: *Provides expert technical and user support to the research community, maintaining the level and quality of services as user numbers grow and providing integrated front-line support covering all DRI pillars.*

Stemming from this exercise, the need for additional detail on the approach to the provision of commercial and community cloud services, and the perceived need to expand on the questions from the 2021 Needs Assessment Survey[8], the Alliance launched a Cloud Survey in late 2022.

Cloud Survey

Alliance's Researcher Council Cloud Survey Steering Group conducted a Cloud Survey in the Winter-Spring 2023. The results of the survey were published in the Alliance Researcher Council Cloud Survey Report[9,10]. **The key findings of the survey** are presented below for reference:

A total of 511 researchers responded to the survey. Quebec (36%) and Ontario (30%) were highly represented, compared to British Columbia (13.7%), Alberta (8.8%) and Nova Scotia (5.1%). The remaining provinces accounted for only ~6% of respondents.

---

[8] Felipe Pérez-Jvostov et al., *Researcher Needs Assessment: summary of what we heard*, (Digital Research Alliance of Canada, 2021), https://alliancecan.ca/sites/default/files/2022-03/needsassessment_alliance_20220126.pdf.

[9] Fares Dhane and Felipe Pérez-Jvostov, *Alliance Researcher Council Cloud Survey Report,* (Digital Research Alliance of Canada, 2023), https://doi.org/10.5281/zenodo.7901152.

[10] Fares Dhane and Felipe Pérez-Jvostov, "alliancecan/Cloud-Survey-Data-Analysis-: First release," Digital Research Alliance of Canada, May, 2023, https://doi.org/10.5281/zenodo.7901271.

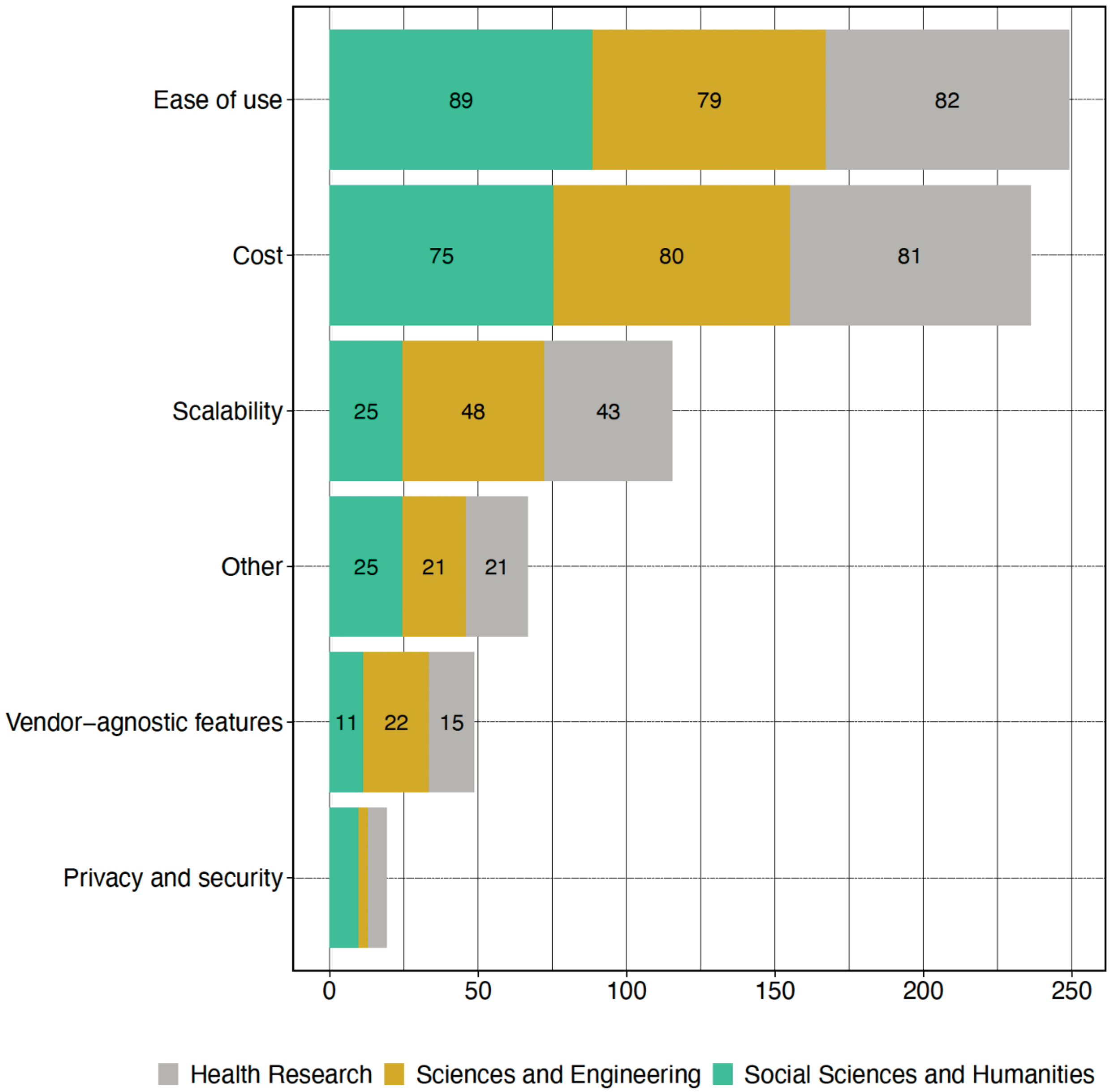

Figure 1: Factors influencing choice of cloud platform

Figure #1 shows that regarding the primary factors researchers considered in choosing which cloud platform to use (Alliance community cloud or commercial) *"...ease-of-use was just as important a factor as cost in choosing cloud services in the Sciences and Engineering (both 79%), and Health Research (82% vs. 81% respectively), and even more important in Social Sciences and Humanities (90% vs. 76%)..."*.

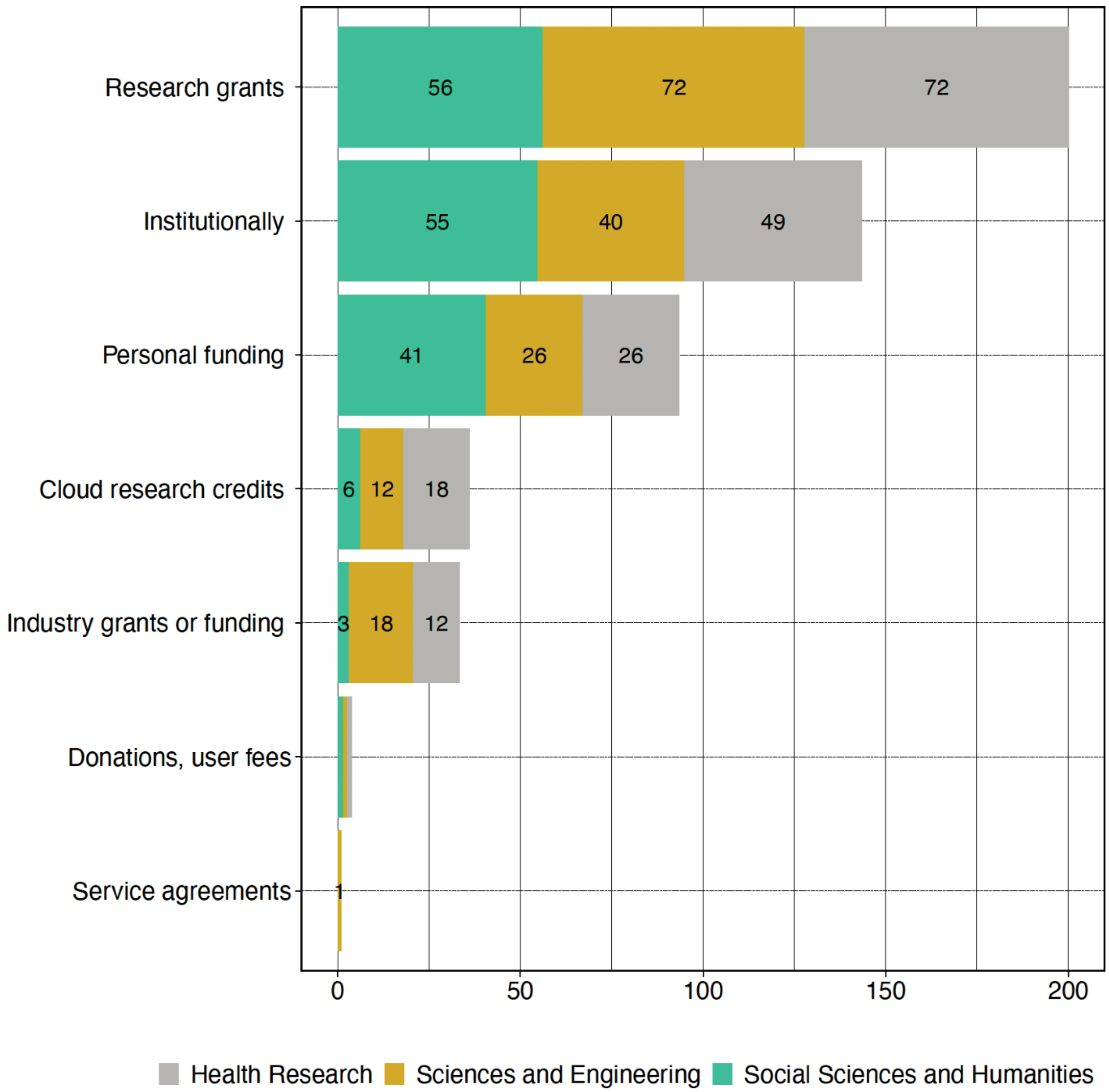

Figure 2: Funds used for commercial cloud.

Figure #2 shows that regarding commercial cloud budget funding *"...close to 60% of respondents acknowledged using commercial cloud (e.g., Microsoft Azure, Google Cloud, Amazon Web Services), covering these expenses through research grants (Social Sciences and Humanities – 56%; Sciences and Engineering – 71%; Health Research – 72%), while also relying on institutional funds (Social Sciences and Humanities – 54%; Sciences and Engineering – 42%; Health Research – 49%). Very few respondents use cloud credits or industry funding to* *cover such expenses. An important observation, however, is that many researchers use their own personal funding to access cloud services, and this is particularly evident in the Social Sciences and Humanities where 40% of respondents identified using their personal funding to cover the costs associated with using commercial cloud – an issue that certainly requires further exploration…".*

The key takeaways from the Alliance Cloud Survey were:

1. *Alliance Community Cloud is not widely used, whether because researchers are using commercial cloud, or they do not know about the Alliance's cloud service, etc. Still, it is essential for those researchers that do–particularly in the Sciences and Engineering.*
2. *Canadian researchers increasingly adopt cloud resources: the number of projects registered to use the Alliance's Community Cloud grew from 220 in 2018 to 950 in 2023.*
3. *Cost and ease of use are the most important decision-making factors for researchers' choice of cloud service.*
4. *Automating and managing software deployment, scaling and management (e.g., Kubernetes and containerization) were identified as service gaps in the Alliance Community Cloud.*
5. *Many researchers use the cloud to store, back up, and share data with collaborators.*
6. *Researchers share privacy and security concerns when using the commercial cloud.*
7. *Researchers use commercial and Alliance Community Cloud to access additional storage and computing.*

The adoption of Cloud technologies is growing in higher education research worldwide, and Canada is no exception. Many researchers across disciplines use the cloud as complimentary compute and storage resources – particularly to back up and share their data with collaborators and peers. While the adoption of commercial cloud is increasing, the Alliance Community Cloud plays an important role for those researchers that do use these services. When asked about the general importance of different cloud services, for example, the Alliance Community Cloud offering was the most important cloud service for researchers in Sciences and Engineering while also playing an important role in Health Research. Researchers also heavily rely on institutional cloud offerings, although it is currently difficult to parse out whether this is a true reliability or lack of clarity between service providers, as both are supported by local support staff. Cost and accessibility play major roles in researchers' decisions. Affordable, and easy-to-use solutions that do not compromise data privacy and security are the most important features for many researchers, particularly because many use institutional or grant funds to cover such expenses.

There is a clear need for increased functionality in the Alliance Community Cloud, including managed Kubernetes, containerization like Docker, and software platforms like JupyterHub. Despite the smaller user base, the Alliance Community Cloud would also benefit from revamped user support, training and documentation. Considering the growing cost of commercial cloud and the broad adoption of cloud technologies, it is conceivable that more researchers will rely on Alliance offerings in the future.

Researcher Council's Cloud Priorities

Alliance's Researcher Council conducted a cloud priorities exercise, publishing the findings in the Spring 2023. **The key findings of the report**[11] were**:**

The priorities identified by the Researcher Council are divided into those of a general nature, those related to the Alliance's community cloud and how the Alliance should provide access to commercial cloud resources.

The Alliance needs to follow the new Cybersecurity Act (Bill C-26). All the vendors and the suppliers of the vendors of both the community cloud and commercial cloud providers should be evaluated through the lens of national security to avoid potential foreign interference and infiltration.

General

1. *Strive to meet the cloud computing requirements of Canadian researchers by providing a refreshed and expanded Alliance Community Cloud that is augmented by the complementary resources of commercial clouds.*
2. *Understand, educate, and provide documentation to the research community on best and safe practices on the use of clouds, and helping researchers ensure the security of their data on both the Alliance Community Cloud and commercial clouds.*
3. *Expand the dedicated cloud support group with a mandate to advise Canadian researchers on cloud technologies (community and commercial) best adapted to their specific needs and provide subsequent technical support, training and documentation.*

Alliance Community Cloud Infrastructure

1. *Ensure that the Alliance Community Cloud provides a wide range of compute, storage and services that are able to adapt to changing demands ("elastic") of the research community. The survey identified a number of services that should be added to the*

[11] Randall Sobie, *Alliance Researcher Council: Meeting the Digital Research Infrastructure Needs of the Canadian Research Community - Update on Cloud Computing,* (Digital Research Alliance of Canada, May 26, 2023), https://zenodo.org/record/7974806.

Alliance Community Cloud (e.g., Docker and Kubernetes) and the need for improved documentation, training material and chat support.

Commercial Cloud Computing

1. *Provide a flexible model for commercial cloud computing that facilitates transparent access or provides in-kind credits to researcher groups to directly purchase cloud computing, storage or services from a vendor of their choice.*
2. *Ensure that Alliance-managed access to commercial clouds is done so that compute, storage and services are provided in a manner that is vendor-agnostic.*
3. *Consider commercial clouds for opportunistic computing when Alliance resources are oversubscribed or offline.*

Problem Statement / Motivation for 2025-30

In 2023 the Alliance launched a second Alliance National Cloud Strategy WG[12], which was asked to focus on the initiatives the Alliance could consider to support researchers in their 2025-30 mandate. This document by the second working group is the result of extensive deliberations and consultations, including those from the first working group's deliberations (as summarized above), the results of the Cloud Survey (as summarized above), recommendations for priorities from the Researcher Council (as summarized above), and approaches taken by the best practices of several international organisations.

The research community needs substantial DRI resources and services, and the Alliance and member institutions currently do not meet either the demand concerning the capacity or the demand for specific software tools and resources.

Cloud Current State

Cloud computing is an approach in which the traditional model of tightly coupled, managed relationships between physical hardware, infrastructure software, and workload software is replaced by a model "enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[13] Current approaches to provisioning cloud resources and services include Infrastructure as a Service (IaaS), Platform-as-a-Service (PaaS), Software as a Service (SaaS),[14] as well as hybrid approaches that involve blending procurement and deployment

---

[12] See Appendix B

[13] Mell and Grance, *The NIST Definition*.

[14] Mell and Grance, *The NIST Definition*.

models, purchasing access to some capabilities from commercial cloud service providers while creating other capabilities themselves through private or community cloud deployments.

Cloud technologies originated as a new model for deploying business enterprise workloads, but in recent years have seen rapid uptake by research communities. The different expectations that research computing users have of their technology resources make their use of cloud platforms markedly different than those of the business technologists who dominated the early stages of cloud technology development. The Alliance's Cloud National Team[15] has observed several cloud technology patterns commonly deployed by researchers on Alliance Community clouds.

Enabling access to cloud-based, container-orchestration workflow systems. Cloud computing provides the necessary virtualization infrastructure to connect artificial intelligence and machine learning researchers to cloud-based, container-orchestrated workflow systems. These play a critical role in managing the complexity of scheduling for huge numbers of individual job repetitions necessary to train machine learning models.

Supporting shared scheduling of HPC-like workloads. Bioinformatics and omics researchers use Community Cloud for projects like CreSCENT and GenAP to interactively schedule HPC-like workloads against centrally held datasets and collaborate on the results through portal interfaces.

Enabling interactive and collaborative workloads. Interactive and collaborative workloads, such as textual and media analysis, are common in the humanities and social sciences. Cloud infrastructure allows portal access to essential tools, such as Voyant and Islandora, which allow researchers in those disciplines to work collaboratively across projects and institutions.

Providing alternative OS tools. Researchers in areas as varied as social sciences, humanities, and geospatial analysis often depend on tools like NVivo or ArcGIS that require specific operating environments - e.g., Microsoft Windows. Community cloud has been a critical support in enabling these environments on demand without being limited to physical workstations maintained by individual projects or institutions.

Supporting rapid access to virtual HPC. Cloud environments can provide HPC capabilities that support key research activities. For example, the made-in-Canada open-source Magic Castle platform provides cloud-hosted virtual HPC clusters that mirror other HPC environments and can be used for training and prototyping.

Offering collaborative portals and publishing. Cloud computing contributes critical access to collaborative analysis portals like GenAP (for bioinformaticians) and Voyant Tools (for textual analyses) and web-specific platforms like Érudit.org (a Canadian scholarly publication platform),

[15] The Alliance Cloud National Team is a strategy-focused group of Alliance Federation cloud technology specialists providing guidance and developing initiatives to extend and improve the Alliance's Community Clouds.

which allows open access to Canadian journal publications. These collaborative platforms provide access to research data and improve visibility and reproducibility in many fields.

Other patterns are suggested by why and how researchers are using the commercial cloud in particular.

1. **Accessing innovative AI tools and algorithms**. AI and related technologies are evolving at a blistering pace, and many of the same large companies that are providing commercial cloud are at the forefront of these developments.
2. **Quick Access**. Researchers can deploy commercial cloud resources simply and quickly without an account approval process or going through a resource allocation process like RAC.
3. **Additional Security.** Commercial cloud vendors provide a variety of services and services for keeping the cloud platform secure and in compliance. These tools can e.g improve visibility and control; allow automation of security tasks, enabling researchers and IT personnel to be more secure by reducing human configuration errors. Since commercial clouds are used by large enterprises, recent privacy and data security standards, and third-party validations are often quickly made available on the platforms.

The current Alliance cloud services are provided in large part by the Arbutus Cloud site hosted at the University of Victoria and augmented by other distributed cloud sites within the Alliance Federation: Béluga Cloud at McGill University/Calcul Québec site located at École de Technologie Supérieure, Graham Cloud at the University of Waterloo, and Cedar Cloud at Simon Fraser University. These sites are operated by Federation partner institutions and attached directly to the joint CANARIE and regional networks forming the National Research and Education Network (NREN) to enable the cost-effective and high-speed transfer of massive research datasets.

A few Canadian consortia also offer community cloud capabilities for academic and research use, including BCNET's EduCloud Server initiative, based on VMWare vSphere and hosted at BCNET facilities in Kamloops and Vancouver, Ontario Council of Research Libraries' Ontario Library Research Cloud based on OpenStack, and Cybera's Rapid Access Cloud, based on OpenStack and hosted at Cybera facilities in Edmonton and Calgary.

In addition to self-hosted and community cloud services, some Canadian researchers are also making use of commercial cloud services, buying access to cloud technologies and capabilities provided by for-profit businesses. It is very difficult to know the exact number of researchers using commercial cloud, but anecdotal evidence suggests it is a substantial proportion of the Canadian community. The Alliance Cloud Connect Pilot (described below) will provide a mechanism to get data on the use of commercial clouds. The commercial cloud market has grown dramatically over the last decade, stratifying significantly as a small group of front-runner cloud service providers have claimed an overwhelming majority of the market share.

According to Hyperion Research, future growth and changes in product and service offerings in the cloud sector will continue to be rapid. They foresee a compound annual growth rate (CAGR) of nearly 17.6% until 2026 for the HPC cloud market compared to a 6.9% CAGR for the on-prem market.[16] A major component in this growth is the movement of workgroup class (defined as systems priced less than $100k USD by Hyperion) workloads to the cloud. Intersect360, on the other hand, forecasts that 2023 was the last year of double-digit growth for the cloud market. Beyond 2024 HPC cloud spending is seen to grow at a general market rate, leveling at roughly 15% of total HPC spending and 23% of total HPC+AI spending.[17] According to Hyperion, within academic institutions, 25% use and plan to continue using commercial cloud for their HPC workloads, while on the other hand 50% do not use and are not planning on using commercial cloud *for HPC workloads* in the next 6-18 months due to structure of academic funding and costs.[18]

The “Tier-One” or “Hyperscale” commercial cloud providers include Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP), which together account for a significant majority of cloud market share. Tier-One cloud providers operate data centers in many locations around the world - each of the largest three offers at least one Canadian location - provide the widest range of cloud technologies from IaaS up through PaaS and SaaS, with profits funding huge investments to produce unique products and capabilities that would be cost-prohibitive for smaller providers to develop. These providers largely rely on the higher profitability of services above IaaS to fund their operations and tend to cost significantly more than lower-tier providers when comparing base IaaS capabilities.

Below tier one, cloud service providers range from those attempting to compete across the whole range of cloud technologies, like IBM Cloud and Oracle Cloud, to those offering lower-cost services focused mostly on IaaS, like OVHcloud, Digital Ocean, Linode, and others. This segment of the market is populated by many cloud service providers, none of which control more than single-digit percentages of overall market share; not all operate their own data center facilities, and not all have access to data centers located in Canada.

The Open Clouds for Research Environments project (OCRE)[19] was funded by the European Commission, granting the GEANT Association to procure commercial cloud services on behalf

---

[16] Earl Joseph et al., *Hyperion Research HPC Market Update (pre SC-22)*, (Hyperion Research, November 2022), https://hyperionresearch.com/wp-content/uploads/2022/11/Hyperion-Research-SC22-HPC-Market_Combined_111022_V2.pdf.

[17] *Worldwide HPC and AI Market - 2022 Actuals, Change from 2021, 2023-27 Forecast*, (Intersect360, May 2023), https://www.intersect360.com/wp-content/uploads/Webinar-Intersect360-WW-HPC-AI-Unified-2022-market-size-and-2023-27-forecast.pdf.

[18] Jaclyn Ludema and Tom Sorensen, *Special Study Results - Cloud Resources are Employed at Nearly Half of HPC Sites and are Expected to Gain Additional Popularity*, (Hyperion Research, May 2023).

[19] “Open Clouds for Research Environments,” August, 2023, https://www.ocre-project.eu/.

of their 40 members: National Research and Education Networks (NRENs) with their national Institutions in 2020. The result was 474 Cloud Framework contracts spread over 40 countries / geographical lots, EU procured through GEANT acting as Central Purchasing Body. Through this Cloud Framework, participating Institutions can procure commercial cloud services from local platforms and resellers of international hyper scalers without having to run their own European tender within the 4 years duration of the framework contract. This approach created an economy of scale, providing participants with beneficial terms, conditions and favourable discounts. The Open Clouds for Research Environments project (OCRE) accelerated cloud adoption in the European research community by bringing together cloud providers, Earth Observation (EO) organizations and the wider European research and education community through validated, ready-to-use service agreements available across the Pan European region. The maximum budget within the OCRE framework is 874M CAD to be procured until the end of 2024. Currently, 26 countries are actively consuming cloud service through the framework and experienced a 45% growth in 2022.

It is interesting to note that the OCRE project saw 1,100 responses to an open EU Tender for commercial cloud, resulting in a portfolio of services with 474 commercial cloud contracts (contracts shall not exceed four (4) years and will be re-procured in 2024) covering 40 European countries with

1. minimum 9 platforms per country, maximum of 17;
2. each platform verified using HNSciCloud workload;
3. all suppliers bound by the GDPR regulations;
4. all platform support for the SAML2/OIDC protocol;
5. favourable T&Cs and discounts negotiated from GÉANT as Central Purchasing Body;
6. ready-to-use contracts without European procurements and no fees for the end-user.

Alliance 2023-25 Commercial Cloud Pilot

The Alliance Cloud Connect Pilot (ACCP) will develop a pilot single “pane of glass” or portal for researchers to access key commercial cloud services, providing access to these services on demand and in a way that simplifies the administrative and technical challenges. The service will simplify researcher’s on-demand access to commercial cloud services to all researchers, facilitate funder and institutional interests in the use/provision of cloud services to the research community. The ACCP will set the stage for the Alliance’s provision of a hybrid cloud service that will not only simplify researcher access to compute, storage, and software, allowing them to focus on research and innovation.

The model the Alliance will use was born out of the U.S. National Science Foundation 2018 award (NSF 2018: Managed Services to Simplify Cloud Access for Computer Science Research and Education). University of California San Diego won the award to develop, in collaboration with San Diego Supercomputer Center and other supercomputing centers, a service platform, named Cloudbank[20]. The development started in 2019 and was announced at the PEARC (Practice & Experience in Advanced Research Computing) in 2021. As of the time of this writing (Spring 2023), Cloudbank has graduated from a pilot into full production, and is serving researchers at several Tier 1 institutions across the United States.

Another important model for the ACCP is Europe's OCRE Project, which focuses on the procurement of commercial cloud services across the EU, for inclusion in a single cloud portal service. The approach that the OCRE team took to reviewing proposals from over 400 multi-national and national CCS providers, and negotiating across multiple jurisdictions and legal contexts will help inform the Alliance's approach for the Pilot and beyond.

Examples and Use Cases

1. Storage
    a. A common use case among Canadian academics using cloud resources is variations on data storage and sharing, commonly e.g., file sharing and collaboration among research team members or global research community in hybrid or geographically dispersed work environments. Such datasets can vary from administrative data, to draft manuscripts, large collections of reference materials, and experimental data, and up to huge public or private datasets available via e.g., Repositories-as-a-Service. Data formats are not restricted to any particular type, and can range from text files to video, audio, images, databases and digitized data to massive genetics or physics datasets.
    b. In some storage use cases a cloud based storage system is leveraged and considered to be a backup of the primary data that is located on-site.
2. Science workflow cloud management portal
    a. Computer Science grad student leveraging commercial cloud via a GUI-based middleware portal that manages images, launching, scaling and budget controls, allowing the student to have on-demand, immediate access to customized compute resources at scale. The stack could consist of e.g., Ronin for cloud management, with persistent head node running Jupyter notebook for prototyping and for running PostgreSQL database instance. Larger simulations would be submitted to the local SLURM scheduler, to be executed in an auto scaling cluster managed by Ronin.
3. Advanced commercial cloud use cases
    a. Data and Analytics
        i. Secure and Compliant Data environments for teams to work both locally and across different research sites.

[20] "Cloudbank," August, 2023, https://www.cloudbank.org/.

    ii. Clean Room and Data Sharing capabilities: Clean Rooms are secure places where analysts can work with confidential data without risking it being seen by unauthorized people. Data Sharing lets authorized people share data securely, keeping sensitive information safe from breaches or unauthorized access. Businesses and organizations can use these capabilities to improve their data security and protect sensitive information.
  b. AI/ML
    i. Access to Foundation Models through a rich set of APIs and platform services.
    ii. Access to pre-trained AI Services provides ready-made intelligence for your applications and workflows.
  c. Applications:
    i. Drug discovery acceleration, via multiple AI training models, allows researchers to increase time to insight. Additionally, the increased performance allows researchers to maximize the utilization of valuable GPU resources, leading to time and potentially cost savings.
    ii. Genomics workflows are high-performance computing workloads. Traditionally, they run on-premises with a collection of scripts. Scientists run and manage these workflows manually, which slows down the product development lifecycle. Scientists spend time administering workflows and handling errors on a day-to-day basis. With Cloud, both commercial and community, the ability to automate and save time from this work focusing on actual research increases.

4. Health research use cases
  a. In early discovery and pre-clinical research, scientists need access to extensive computing power to perform tasks such as computational chemistry simulations or large-scale genomic analyses. However, the procurement, setup, and management of a dedicated on-premises high-performance computing (HPC) cluster, or application process and wait times for substantial compute resources, can be costly and take months.
  b. The Tier 1 commercial cloud providers provide access to infrastructure without upfront investments, including powerful HPC environments that can quickly scale up during peak demand periods, and scale back down again when demand has decreased. Orchestrating services can simplify the use and management of these HPC environments, making them more accessible to everyday users.
  c. Managed personal health information (PHI) compliant cloud based storage for health data, allowing linking of arbitrarily large public and sensitive datasets for health research.[21]

[21] Richard F. Wintle et al., *PHI-Compliant Computing and Storage: a critical need for Canadian biomedical and health Research*, (New Digital Research Infrastructure Organization, December, 2020), https://alliancecan.ca/sites/default/files/2022-03/ndrio-white-paper-sickkids-tcag.pdf.

    d. The iReceptor Data Integration Platform (iReceptor) is an implementation of the adaptive immune receptor repertoire (AIRR) Data Commons envisioned by the AIRR Community. iReceptor is a Distributed Data Management system and Scientific Gateway for mining and analyzing "Next Generation" sequence data from immune responses, consisting of Arbutus community cloud based AIRR-seq data set repositories and a web portal. iReceptor lowers the barrier to immune genetics researchers who need to federate large, distributed, AIRR-seq data sets to answer complex questions about the immune response.[22]

5. Elastic Compute
    a. Most people, when thinking of cloud computing, think of the ease with which they can procure resources when needed. This is only one aspect of elasticity. The other aspect is to contract when they no longer need resources. Scale-out and scale in. Scale up and scale down. The elastic nature of commercial cloud allows users to quickly adapt to changes in demand, easily adding or removing instances as necessary and paying only for the resources they consume, making it a cost-effective and efficient solution for hosting applications, running workloads, and performing various computational tasks in the cloud.
6. Complex system use case / Digital Twins
    a. Digital twins are virtual replicas or models of physical objects, processes, or systems that are created using real-time data and simulation techniques. They provide a means to emulate, monitor, analyze, and optimize the performance and behaviour of their real-world counterparts. Digital twins are becoming more common in research (they have been used in industry for some time) as a way to combine data from sensors, IoT devices, and other sources with advanced analytics and modelling. By simulating and visualizing real-world scenarios, digital twins allow for the testing of different scenarios in a research context and facilitate changes in the physical world.
7. Quantum research & development
    a. Quantum computing market is in a very strong research & development and growth phase until the foreseeable future with multiple, different, and very expensive competing technologies and systems in development. Accessing a variety of quantum computing platforms offered by commercial cloud providers will enable Canadian researchers to test and prepare for emerging quantum computer technologies.
8. Canadian Advanced Network for Astronomical Research (CANFAR)
    a. CANFAR's Science Portal platform allows Canadian and international astrophysicists to store, share, and calibrate and analyze very large observational datasets. The platform utilizes customized containers and Kubernetes to provide virtual machines (VMs) with specialized software installed, running on Alliance Cloud.

---

[22] Felix Breden and Brian Corrie, *iReceptor – A case study in the challenges/opportunities in Canadian DRI,* (New Digital Research Infrastructure Organization, December, 2020), https://alliancecan.ca/sites/default/files/2022-03/ireceptordriwhitepaper-final.pdf.

9. Cyber Forensic Research
    a. Cyber forensic research involves systematically investigating and analyzing digital evidence to uncover, understand, and prevent cybercrime or malicious activities. It encompasses a multidisciplinary approach that combines computer science, law enforcement techniques, and investigative methodologies to collect, preserve, and analyze digital artifacts such as log files, network traffic, computer systems, and storage devices. The research aims to uncover the who, what, when, where, and how of cyber incidents, helping organizations mitigate threats, improve security measures, and enhance cyber resilience.
10. Multi-Jurisdictional / Global
    a. Algonquian Dictionaries Project and Algonquian Linguistic Atlas digital humanities project at Carleton University documents languages, creates dictionaries and other language learning and teaching tools in collaboration with Indigenous communities, organisations and universities across provinces. The research infrastructure is a cloud based customised open-source data management platform that runs in a Linux Apache MySQL PHP (LAMP) Wordpress web server environment.[23]
11. Selected European Union's Open Clouds for Research Environments (OCRE) use cases[24]
    a. A prototypical semantic Earth Observation (EO) data cube covers the entirety of Austria and contains all Sentinel-2 images available since the start of Sentinel-2A in 2015. The cloud-based platform allows users/non-EO experts to analyse terabytes of Sentinel-2 data within the browser portal without the need of any programming skills and integrate results GIS-ready into their workflows. The platform is implemented using open source and state-of-the-art scalable cloud-based infrastructure with lightweight virtualisation using containers (Docker).
    b. Leveraging cloud resources to optimise drug delivery to lungs -project simulated drug inhalation using realistic dosages leveraging computational particle-fluid dynamics and multiscale physics modelling. The commercial cloud access provided extensive computational resources with implementing various HPC configurations such as across more CPUs than previously with more available RAM that are necessary to perform the first ever simulations of realistic aerosol drug deposition in lung airways.

---

[23] Marie-Odile Junker and Delasie Torkornoo*, Indigenous language technologies and online resources: Algonquian Dictionaries Project and Algonquian Linguistic Atlas,* (New Digital Research Infrastructure Organization, December 2020), https://alliancecan.ca/sites/default/files/2022-03/junker_torkornoo_white_paper_ndrio_2020.pdf.

[24] "OCRE Success Stories," Open Clouds for Research Environments, June, 2023, https://www.ocre-project.eu/success-story.

Challenges and Opportunities

(Note: Each point in this section is intended to address challenges and opportunities for both commercial and community cloud)

1. Talent, skills and staffing

In order to leverage (IaaS, and PaaS in particular) cloud resources, research project teams must develop capabilities in areas including system and network administration, storage and backup, infrastructure design, and cloud architecture. Each skill area is specialized and can be hard to recruit/obtain, and retain. There is also the increasing importance of cybersecurity and data privacy, which adds additional complexity and responsibilities to these roles. Hyperion Research reports that 'lack of knowledge, or skilled cloud computing support staff[25] is the biggest barrier to expanding use of HPC clouds in academia.

2. The difficulty of using traditional HPC

Traditional cluster-based HPC resources are an alien concept to most users, because they use a command line interface and involve potential complex hardware and software architectures, myriad policies and a sophisticated queuing/scheduler system. This barrier is somewhat lowered by the implementation of web portals such as JupyterHub on Alliance's community cloud. Cloud resources can come with thousands of scientific software packages pre-installed, and the barrier of entry can be low when offered as software-as-a-service (SaaS). In contrast, current community cloud offering in the Alliance is mostly infrastructure-as-a-service (IaaS), and is much harder to use, as researchers must be system administrators to deploy and configure their virtual machines. Improving the ease-of-use by leveraging higher abstraction level is a key opportunity, and a challenge, both for the community cloud offering and the HPC resources of the Alliance.

3. Adapting computationally intensive, persistent workflows and shared scheduling

Some compute-intensive jobs could run equally well in a cloud environment, particularly since HPC cluster environments can be nearly replicated in the cloud. For example, the Alliance's Magic Castle project could potentially form a basis for HPC workloads to be performed in the cloud with a nearly identical environment to the current HPC clusters. From the researcher's perspective it could look very much like logging on to any other HPC cluster. This is of particular interest for bursty use cases, and for cases in which an environment shared with thousands of users is not appropriate, such as research that has higher than average privacy concerns, or for

[25] Joseph, *Hyperion Research HPC Market Update.*

teaching. However, it is important to note that the hardware differences between systems imply that certain types of jobs, for example jobs that require low latency interconnects such as distributed memory jobs, or jobs that are I/O limited, would be better suited to more traditional HPC clusters.

4. Accessibility in terms of complexity, as well as cost, and cost management

Multiple consultations with the research community indicate that cloud technology is in high demand and is not being utilized to its full potential. The barriers experienced vary depending on discipline, institution and use-case, but some prevalent barriers are: cost, knowledge and understanding of the technology, user experience of the technology, licensing challenges, and unique software requirements. For example, research funding models may not be conducive to supporting these types of needs, and individual researchers may have much less exposure and support to the technology within their research communities. Similarly, smaller institutions and underrepresented research groups may experience inequitable resourcing and access to funding and technical support required to fully utilize cloud technology in research. Additionally, properly accounting for and adhering to all legal, contracting, and procurement requirements can be challenging when leveraging commercial clouds, particularly since the ease-of-use via “simply enter your credit card number and click OK” use model of commercial clouds potentially hides these complexities from the end-users.

A detailed cost-effectiveness assessment of community cloud versus commercial cloud will depend critically on the services required by the researchers. The Alliance needs to develop the expertise to identify when a commercial cloud solution will be more cost- or time-effective and well managed than a community cloud solution or a HPC cluster solution. The Alliance should:

A. Establish a framework for estimating and comparing total-cost-of-ownership (TCO) for common cloud use cases and services in the community cloud v. commercial cloud.
B. Work with commercial cloud vendors to unify the acquisition and purchase of cloud services to incentivize improvements, to provide guaranteed service delivery, to obtain better and longer-term pricing, and to work with other funding agencies to better enable the use of grant funds towards commercial cloud expenses.
C. Establish robust cloud cost management and budgeting middleware / portals.
D. Investigate potential for allocation of additional funds to allow researchers to access commercial cloud services.
E. Establish effective research data management in the cloud.

Cloud providers offer a variety of research data management (RDM) services (e.g., metadata description, data curation, preservation, etc.). For example, large commercial cloud providers have AI-driven tools such as text extraction and analysis, which can simplify and enhance metadata extraction. Similar tools are available across all stages of the RDM lifecycle, and tailored to specific domains. The opportunity is how to leverage these tools regardless of the cloud service/platform in which the data resides - in essence the provision of a platform-agnostic services approach to domain workflows. There are also cloud-based platforms with a focus on RDM and Open Science workflows, such as the Open Science Framework, Dryad and Figshare.

Another key RDM consideration is data storage and data location (in the light of jurisdictional limitations and regulations), and especially long-term, or archival storage and ingress and egress related technical considerations and cost concerns. Developing a consistent approach, ideally one based on a policy-based storage approach, is a key component of the Alliance's strategy.

5. Facilitating best practices in the use of cloud and associated SW

In order to deliver on the promise of cloud, there is a need for a properly resourced Cloud Researcher Support National Team to support, document, and promote best practices in cloud adoption, and provide more systems administrative support and services for cloud for users. This team should also closely work with the Research Support Team and discipline specific support teams. Support that would be useful going forward would include the following:

A. Technical writers to support detailed user documentation.
B. Administrative support for communications, cloud usage data statistics, feedback mechanisms.
C. Incorporating best practices for cloud usage into the national training strategy. Consideration of this should be reflected in resources for administrative, technological and instruction support, and should incorporate different needs for different delivery models, accessibility concerns, and be co-developed with the appropriate stakeholders (e.g., cloud stakeholders, cybersecurity, training stakeholders, research support).

6. The resiliency of cloud resources

Strategic alignment and awareness of a rapidly-evolving threat landscape are required for cloud computing resilience, which ensures that research resources and services remain accessible, and that researchers' data and development efforts will be protected and recoverable.

The increasing prevalence of cyberattacks in the form of phishing and ransomware has increased the pressure on technology departments to prepare for potential cyberattacks against their organization. As a result, organizations are turning to Disaster Recovery as a Service

(DRaaS) coupled with Backup as a Service (BaaS) to guarantee recoverability following such an attack to their stakeholders.

The current state of DRI reflects a design based on cost performance, thus resiliency requirements are typically not a priority. The Alliance has an opportunity to create an approach to extend to other means of data access and resiliency by using commercial cloud services when in need and required.

The Alliance should, in collaboration with commercial cloud service providers, conduct a needs assessment and costing estimates for creating a comprehensive business continuity plan embedded with robust Research Data Management services designed to protect proprietary data produced by researchers, programs and projects. This should be further supported by efforts to increase awareness of data sovereignty, data security, and ongoing investments in current infrastructure to maintain and support researchers throughout the research lifecycle.

7. Leveraging cutting-edge technologies and software tools

The sheer scale and financial resources of commercial cloud service providers allow for a very wide service portfolio, serving a large roster of use cases. Such services include sophisticated PaaS and SaaS offerings, and access to leading edge hardware technologies, e.g., custom AI and ML accelerator chips for accelerated production runs, and quantum computing systems for testing and prototyping different quantum technologies. While providing access to a large portfolio of commercial cloud services, the Alliance needs to carefully assess when a particular service or research workflow should be built in-house and offered in the community cloud, or when the best return-on-investment is achieved by leveraging commercial clouds.

8. Risk of inequity

Smaller projects, projects based at small institutions, and projects developed in disciplines with incompatible funding schemes and/or less technical training and resources, are all at risk for inequitable access to cloud technologies. Given the particular affinity of cloud to innovative projects and novel research, this inequity could (and arguably already has) stunt the potential for new and developing compute-intensive research across Canada.

9. Traditional CAPEX/OPEX divide

Commercial cloud expenses are usually considered operational expenses, providing a challenge for traditional DRI funding models that are often based on building on-premise infrastructures via capital expenditure funding. If the CAPEX/OPEX ratio in available funding is explicitly restricted, such restrictions can result in either no funding at all, or in suboptimal solutions (e.g., if an on-prem solution is built even if a commercial cloud solution would have been available and would have offered better ROI). DRI services funding should adopt a flexible approach regarding CAPEX/OPEX mix, and require ROI-type assessment and justification from the grant applicants regarding optimal service delivery system, community cloud v. commercial cloud.

10. Improved security

Cloud environments can provide opportunities for improved security through a shared responsibility between the user and cloud provider, or in some cases, the full assumption of some aspects of the security envelope by the cloud provider. For example in PaaS and SaaS environments the cloud provider would be responsible for automatic software updates, removing the responsibility from individual researchers and research groups. Additionally, in mature PaaS and SaaS environments the cloud provider can build and provide the complex custom hardware, software, and workflow stacks that are built and adhere to specific security and policy standards, e.g., for the purposes of storing and processing sensitive data, while keeping in mind data sovereignty requirements.

Vision, Goals and Recommendations

Vision

Leverage an enhanced and rich suite of cloud services to accelerate research projects where applicable and develop a long-term strategy for using cloud computing to increase research capacity and knowledge creation.

Key Goals

1. Meet the DRI resource demand from the Canadian academic research community leveraging traditional HPC, community cloud, and commercial cloud resources, combined with effective and efficient use of hybrid cloud services.
2. Provide an on-demand approach to the provision of cloud resources and one tailored to the needs of specific communities of practice.

Recommendations

General

1. **Build a hybrid cloud service and a suite of associated services**, including appropriate research software and storage.
2. **Leverage federated identity services**, including granular information about researchers, institutional agreements and partnerships, and international frameworks (e.g., AAI) to facilitate an on-demand cloud resource delivery that is scalable and cost-effective.
    a. **Ensure collaboration with the EU's AAI framework** and adoption of best practices from the EU collaboration.
    b. **Ensure support for and integration with regional, national and international privacy legislation** and specific digital identity frameworks.

People and Community

3. **Ensure researchers' needs for specific research software and tools are met**, regardless of which computing platform they are using.
    a. **Adopt a *marketplace* approach to facilitating access to applications**, allowing institutional, regional, national and international participation.
    b. **Consider a research version of the CANARIE DAIR platforms**, providing pre-packaged instances of software, data and services.
    c. **Build a dedicated team of cloud researcher support specialists** to help inform and guide research groups on the best cloud services and practices for their workloads.
    d. **Support creation of new cloud tools and platforms** for the Canadian research community, aligned with Alliance Research Software Strategy's substantial recommendations supporting new and existing software initiatives.

Policy

4. **Ensure support for Open Science approaches in the use of the cloud** as the dominant emerging approach to conducting research.
    a. **Facilitate integration of open science approaches**, such as easy deposit of data and software from cloud platforms and services.
    b. **Include cloud platforms and tools in a national registry of DRI tools and services**, including appropriate PIDs to reflect open science best practices.
5. **Ensure support for and integration with indigenous DRI resources and services**, including respecting indigenous data sovereignty.

Infrastructure and Services

6. **Build on the 2023-25 enhancement to the Alliance's Community Cloud capacity and services.**
    a. **Commit to an annual renewal procurement process**, ensuring reliable availability of hardware and growth to meet expanding researcher needs.
    b. **Leverage the trends in resource allocation (RAC) and rapid access service (RAS)** to predict the future community cloud CPU, GPU, and storage needs.
    c. **Maintain sufficient operational overhead to ensure researchers have immediate access to resources**; due to the on-demand nature of cloud computing, the capacity should be 20% beyond the expected needs to allow for bursty workloads.
    d. **Develop new services to enhance the functionality of community clouds**, based on strategic priorities and an ongoing needs assessment, including capacity expansion based on RAC/RPP requests and other anticipated demands.
7. **Review the outcomes of the 2023-25 Commercial Cloud Pilot, Alliance CloudConnect, and detail additional requirements and capacity for 2025-30**.
    a. A key goal of the 2023-25 Cloud Pilot is to derive greater detail on commercial cloud costs, usage, software use, and support options. These details should inform the specific initiative(s) for the 2025-30 mandate.
    b. The framework for tracking these elements of cloud should form the basis of an annual review for 2025-30, and beyond.
8. **Provide selected mature cross-discipline cloud-based platforms and workflows as fully funded and supported national services;** such gateways and workflows should be discipline agnostic and widely used, and already in production or nearly production ready (e.g., Syzygy, MagicCastle, etc.).
9. **Ensure quality network connectivity, to both** National Research and Education Network and global Internet, **for both commercial and community cloud**.
10. **Ensure a national approach to sufficient data storage and transfer** and one that is in sync with the Alliance's HPC Strategy for 2025-30.

Data

11. **Ensure an appropriate approach to data movement, storage, use, and in accordance with appropriate legislative and regulatory frameworks**.
    a. **Facilitate an appropriate approach to the movement of data between community and commercial cloud contexts**, including the ingress and egress in the commercial cloud context. Leverage the opportunity to work with commercial cloud providers' policies and approaches to this one.
    b. **Ensure attention to data sovereignty in the provision of cloud services**, including working with Canadian cloud providers, and in sync with the Alliance's HPC Strategy for 2025-30.

Security

12. **Ensure that best practice security policies and approaches are part of the community and commercial cloud services**.
    a. In alignment with Alliance Federation's Cybersecurity Strategy and Roadmap. For further details, see Appendix B.
    b. Delineate roles and responsibilities in this and other contexts (actors, institutions, etc).
13. **Ensure appropriate support and clearance for all appropriate levels of security** clearance and adherence for higher education and government research.

Training & Support

14. **Develop an approach to training for all cloud services offered and supported** by the Alliance and stakeholders, and one that is in sync with the Alliance's National Training Strategy for 2025-30.
    a. **Ensure a collaborative approach to training**, both for general use and domain-specific use and in sync with the international deployment and use of common commercial cloud services.
15. **Develop an approach to delivering effective support at all levels and for all types of cloud**.
    a. **Develop an approach and service details for the Cloud National Team** and one that includes community and commercial cloud.
16. **Increase the capability of training and support of the use of RS in the cloud.** Ensure that training and support for emerging Platform-as-a-Service offerings from the ARC/Cloud team are developed (content) and resourced (FTEs) appropriately. Provide tools, expertise and support for code development and management including Continuous Integration and Continuous Deployment (CI/CD) pipelines.

Governance, Coordination, & Partnership

17. **Ensure the development of sustainable/innovative/compliance approaches to cloud procurement**, including international partnerships.
    a. Determine how best to achieve the removal of "artificial" CAPEX/OPEX distinction in DRI funding.
18. **Ensure national and provincial funding agencies are part of a process** that facilitates planning and use of the cloud for research purposes.
    a. **Work with funding agencies on the development of their policies** for data security.
    b. **Work with funding agencies to ensure a proactive understanding of the potential need for cloud resources**, starting with the various funding call submissions.
    c. **Work with federal agencies on the integration of machine-actionable DMPs** (maDMPs) into submission pipelines to provide detail on proposed DRI demand.

Appendix A - Working Group Members

2022 - 1st Alliance National Cloud Strategy WG Members

Name	Organisation
Joseph Abramo	The Alliance
Roy Chartier	The Alliance
Mark Leggott	The Alliance
Qian Zhang	The Alliance
Jacob Boschee	University of British Columbia
Jeff Albert	University of Victoria
Victor Ionescu	Calcul Quebec
Chris Geroux	ACENET
Pekka Sinervo	University of Toronto/Researcher Council Member
Emmanuel Chateau-Dutier	L'Université de Montréal/Researcher Council Member
Pieter Botha	University of Guelph
Deborah Stacey	University of Guelph
Lydia Vermeyden	The Alliance

2023 - 2nd Alliance National Cloud Strategy WG Members

Name	Organisation
Mark Leggott (Chair)	The Alliance
Matthew Smith (Member)	University of British Columbia
Ryan Enge (Member)	University of Victoria
Liseanne Cadieux (Member)	The Alliance
Rebecca Davis (Member)	Researcher Council/University of Manitoba
Mark Wolff (Member)	CANARIE
Terry Peckham (Member)	Researcher Council/Saskatchewan Polytechnic
Danny D'Amours (Member)	Canadian Government Representative/National Research Council (NRC)
Shava Smallen (SME)	SDSC, Sherlock Cloud
Monique Pellinkhof (SME)	GÉANT/OCRE
Morgan Lim (SME)	Commercial Cloud/Amazon
Diego Magalhães (SME)	Commercial Cloud/Amazon
Frank Currie (SME)	Commercial Cloud/Google
John Weigelt (SME)	Commercial Cloud/Microsoft
Sarah Finney (SME)	Small Canadian Commercial Cloud Representative/Think On

Kimberley Hartley (Support)	The Alliance
Fares Dhane (Support)	The Alliance
Seppo Sahrakorpi (Support)	The Alliance

Appendix B - Cybersecurity in the Cloud

Alliance Federation's Cybersecurity Strategy and Roadmap recommends awareness and training on data security protocols including Identification and Protection of sensitive content. In addition, the Cybersecurity document conducted an assessment as it relates to data security for both in-house (i.e., community cloud) and commercial cloud. It should also be noted that data security in commercial cloud providers is a shared responsibility as they do not *'assess themselves against standard Government of Canada (GC) guidelines, such as ITSG-33. Instead, they meet internationally recognized certifications, such as the International Organization for Standardization's ISO 27001. The GC has published the cross-references of ITSG-33 with prevalent cloud industry certifications. The GC has also published the Direction on the Secure Use of Commercial Cloud Services: Security Policy Implementation Notice (SPIN)*[26] *to support departments and agencies in their understanding of existing Treasury Board of Canada Secretariat (TBS) security policy requirements in the context of cloud computing*[6]*'.* As the Alliance increases adoption of commercial cloud services, it has its share of responsibility as it relates to data security, this further supports the importance of training. Another key element to data security in the cloud could potentially be to put in place a mechanism in the form of a short questionnaire that helps with determining the security classification of content. The questionnaire could be built into the onboarding process so that once the security classification of the data is known, the data could automatically be deposited to a proper secure platform.

[26] "Direction on the Secure Use of Commercial Cloud Services: Security Policy Implementation Notice (SPIN)," Government of Canada, June, 2022, https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/cloud-services/direction-secure-use-commercial-cloud-services-spin.html.